TÉCNICO LISBOA

Aprendizagem 2021/22

Homework II – Group 072

I. Pen-and-paper

1)

Basis Function $\phi_j(x) = \|w\|_2^j$

Considerando cada linha como uma observação (excluindo o output) ficamos com 8 observações com D=3

$x_1 = [1\ 1\ 0]$, $x_2 = [1\ 1\ 5]$, $x_3 = [0\ 2\ 4]$, $x_4 = [1\ 2\ 3]$

$x_5 = [2\ 0\ 7]$, $x_6 = [1\ 1\ 7]$, $x_7 = [2\ 0\ 2]$, $x_8 = [0\ 2\ 9]$

Obtemos uma regressão polinomial com a equação:

$$f(x,w) = \sum_{j=0}^{3} w_j \cdot \phi_j(x) = w_0 + w_1\|x\|_2 + w_2\|x\|_2^2 + w_3\|x\|_2^3$$

Para achar os pesos representamos a amostra numa design matrix

$$\Phi = \begin{bmatrix} 1 & \sqrt{2} & \sqrt{2}^2 & \sqrt{2}^3 \\ 1 & 3\sqrt{3} & (3\sqrt{3})^2 & (3\sqrt{3})^3 \\ 1 & 2\sqrt{5} & (2\sqrt{5})^2 & (2\sqrt{5})^3 \\ 1 & \sqrt{14} & (\sqrt{14})^2 & (\sqrt{14})^3 \\ 1 & \sqrt{53} & (\sqrt{53})^2 & (\sqrt{53})^3 \\ 1 & \sqrt{3} & (\sqrt{3})^2 & (\sqrt{3})^3 \\ 1 & 2\sqrt{2} & (2\sqrt{2})^2 & (2\sqrt{2})^3 \\ 1 & \sqrt{85} & \sqrt{85}^2 & (\sqrt{85})^3 \end{bmatrix}$$

em que

$\|[1\ 1\ 0]\| = \sqrt{2}$

$\|[1\ 1\ 5]\| = 3\sqrt{3}$

$\|[0\ 2\ 4]\| = 2\sqrt{5}$

$\|[1\ 2\ 3]\| = \sqrt{14}$

$\|[2\ 0\ 7]\| = \sqrt{53}$

$\|[1\ 1\ 7]\| = \sqrt{3}$

$\|[2\ 0\ 2]\| = 2\sqrt{2}$

$\|[0\ 2\ 9]\| = \sqrt{85}$

$$t = \begin{bmatrix} 3 \\ 2 \\ 8 \\ 4 \\ 5 \\ 3 \\ 7 \end{bmatrix}$$

Onde a solução fechada é:

$$w = (\Phi^T \cdot \Phi)^{-1} \cdot \Phi^T \cdot t =$$

$$w = \begin{bmatrix} 4,583527221 \\ -1,687204806 \\ 0,337737254 \\ -0,073306742 57 \end{bmatrix}$$

ficando assim com a equação que representa a regressão polinomial:

$$f(x,w) = 4,583527 - 1,687204\|x\|_2 + 0,337737\|x\|_2^2 - 0,073306\|x\|_2^3$$

TÉCNICO LISBOA

Aprendizagem 2021/22

Homework II – Group 072

2)

Consideremos 2 observações de teste

$x_9 = [2\ 0\ 0]$ e $x_{10} = [-1\ 2\ 7]$

Para acharmos a RMSE temos que prever os valores de cada observação no modelo de regressão polinomial. Para isso achamos o valor de x_9 e x_{10} através da função descreve a regressão polinomial

$\|x_9\|_2 = \|[2\ 0\ 0]\|_2 = 2$

$\|x_{10}\|_2 = \|[-1\ 2\ 7]\|_2 = \sqrt{6}$

$$F(x_9, w) = 4{,}583521 - 7{,}6872048 \times 2 + 0{,}3377373 \times 2^2 - 0{,}07330674 \times 2^3 = 2{,}45367$$

$$F(x_{10}, w) = 4{,}583521 - 7{,}6872048 \times \sqrt{6} + 0{,}3377373 \times \sqrt{6}^2 - 0{,}07330674 \times \sqrt{6}^3 = 2{,}28759$$

Dado os valores anteriores, e que temos 2 observações sabemos que:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} = \sqrt{\frac{(2{,}45367 - 2)^2 + (2{,}28759 - 4)^2}{2}} \simeq$$

$$\simeq 1{,}25672$$

TÉCNICO LISBOA

Aprendizagem 2021/22

Homework II – Group 072

3)

Fazendo uma equal depth-binarization em y_3 e output, ficamos com:

	y_1	y_2	y_3	output
x_1	1	1	N	N
x_2	1	1	P	N
x_3	0	2	P	N
x_4	1	2	N	N
x_5	2	0	P	P
x_6	1	1	N	P
x_7	2	0	N	P
x_8	0	2	P	P

output $\Leftrightarrow$ O

$$E_{start} = E\left(\frac{4}{8}, \frac{4}{8}\right) = -\left(\frac{4}{8}\log_2\left(\frac{4}{8}\right) + \frac{4}{8}\log_2\left(\frac{4}{8}\right)\right) = 1$$

$y_1 = 0$ ↓ $\#\{O = N\} = 1$, $\#\{O = P\} = 1$

$y_1 = 1$ ↓ $\#\{O = N\} = 3$, $\#\{O = P\} = 1$

$y_1 = 2$ ↓ $\#\{O = N\} = 0$, $\#\{O = P\} = 2$

$$E(y_1 = 0) = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = 1$$

$$E(y_1 = 1) = -\frac{1}{4}\log_2\left(\frac{1}{4}\right) - \frac{3}{4}\log_2\left(\frac{3}{4}\right) \simeq 0,8113$$

$$E(y_1 = 2) = -\frac{2}{2}\log_2\left(\frac{2}{2}\right) = 0$$

$$E_{y_1} = \frac{2}{8} \times E(y_1 = 0) + \frac{4}{8} E(y_1 = 1) + \frac{2}{8} E(y_1 = 2) =$$

$$= \frac{2}{8} \times 1 + \frac{4}{8} \times 0,8113 + \frac{2}{8} \times 0 = 0,65565$$

TÉCNICO LISBOA

Aprendizagem 2021/22

Homework II – Group 072

$y_2=0$ | $y_2=1$ | $y_2=2$

$\#30=N\,y=0$ | $\#30=N\,y=2$ | $\#30=N\,y=2$

$\#30=P\,y=2$ | $\#30=P\,y=1$ | $\#30=P\,y=1$

$E(y_2=0)=-\frac{2}{2}\log_2\left(\frac{2}{2}\right)=0$

$E(y_2=1)=-\frac{1}{3}\log_2\left(\frac{1}{3}\right)-\frac{2}{3}\log_2\left(\frac{2}{3}\right)\simeq 0,9183$

$E(y_2=2)=-\frac{1}{3}\log_2\left(\frac{1}{3}\right)-\frac{2}{3}\log_2\left(\frac{2}{3}\right)\simeq 0,9183$

$E_{y_2}=\frac{2}{8}\times E(y_2=0)+\frac{3}{8}\times E(y_2=1)+\frac{3}{8}\times E(y_2=2)=$

$=\frac{2}{8}\times 0+\frac{3}{8}\times 0,9183+\frac{3}{8}\times 0,9183=$

$=0,688725$

$IG(y_2)=$ Estart $-E(y_2)=1-0,688725=0,311275$

$y_3=N$ | $y_3=P$

$\#30=N\,y=2$ | $\#30=N\,y=2$

$\#30=P\,y=2$ | $\#30=P\,y=2$

$E(y_3=N)=-\frac{2}{4}\log_2\left(\frac{2}{4}\right)-\frac{2}{4}\log_2\left(\frac{2}{4}\right)=1$

$E(y_3=P)=-\frac{2}{4}\log_2\left(\frac{2}{4}\right)-\frac{2}{4}\log_2\left(\frac{2}{4}\right)=1$

$E(y_3)=\frac{4}{8}E(y_3=N)+\frac{4}{8}E(y_3=P)=\frac{4}{8}\times 2=1$

$IG(y_3)=$ Estart $-E(y_3)=1-1=0$

Com isto concluímos que o maior IG é o de y_1 entao y_1 é o nó da árvore e ficamos com:

$y_1=0$

y_2	y_3	output
2	P	N
2	P	P

$y_1=1$

y_2	y_3	output
1	N	N
1	P	N
2	N	N
1	N	P

$y_1=2$

y_2	y_3	output
0	P	P
0	N	P

TÉCNICO LISBOA

Aprendizagem 2021/22

Homework II – Group 072

No entanto, as partições ($y_1=0$ e $y_1=1$) ainda têm incerteza, então temos que repetir o processo para ambas

$$E(\text{start}|y_1=0) = -\frac{1}{2}\times\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1$$

$y_2=2$	$y_3=P$
$\#_{y_0=N}=1$	$\#_{y_0=N}=1$
$\#_{y_0=P}=1$	$\#_{y_0=P}=1$

$$E(y_1=0 \wedge y_2=2) = -\frac{1}{2}\log_2\left(\frac{1}{2}\right)$$

$$IG(y_1=0 \wedge y_2=2) = 1-0,5 = 0,5$$

$$E(y_1=0 \wedge y_2=2 \wedge y_3=P) = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) = 0,5$$

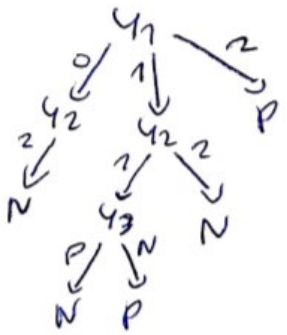
$$IG(y_1=0 \wedge y_2=2 \wedge y_3=P) = 1-0,5 = 0,5$$

Dado que o IG de ambos é igual escolhemos um qualquer (neste caso que se $y_2=2 \wedge y_3=P$ então o output é N).

Também pela tabela acima podemos concluir que se $y_1=1 \wedge y_2=2$ então $O=N$ e que se $y_1=1 \wedge y_2=1 \wedge y_3=P$ então $O=N$.

Pelo caso referido em cima como o IG se $y_1=1 \wedge y_2=1 \wedge y_3=N$ ser o mesmo escolhe-se que nesse caso o output é P

Assim sendo, ficamos com a árvore



Aprendizagem 2021/22

Homework II – Group 072

4)

Como podemos ver pela RMSE e a accuracy, o modelo não preve bem os dados de teste devido á RMSE ser alta e a accuracy baixa. Porém, podemos ver que para X9 o valor previsto é muito perto do atual (output), contudo, também podemos ver que o X10 é um noise point e que afeta drásticamente a previsão do modelo, dado que o valor previsto para X10 (2,28159) é muito menor que o valor do output (4).

$$\text{accuracy} = \frac{\text{nº previsões corretas}}{\text{total número de previsões}} = \frac{1}{2} = 0{,}5$$

II. Programming and critical analysis

5) Answer 5

6) Answer 6

7) Answer 7

III. APPENDIX

END